AMENDMENT

IN THE SPECIFICATION

Please replace the paragraph on page 2, lines 13-25 with the following:

Identification of target loci and the isolation of associated genes using molecular markers has been reported (Liu *et al., Proc. Natl. Acad. Sci. USA, 96*:6535-6540 (1999); Muramoto *et al., The Plant Cell, 11*:335-347 (1999); Bowman and Smyth, *Development, 126*:2387-2396 (1999); Michaels and Amasino, *The Plant Cell, 11*:949-956 (1999); Ha *et al., The Plant Cell, 11*:1153-1163 (1999); Walker *et al., The Plant Cell, 11*:1337-1349 (1999); Sedbrook *et al., Proc. Natl. Acad. Sci. USA, 96*:1140-1145 (1999); Kiyosue *et al., Proc. Natl. Acad. Sci. USA, 96*:4186-4191 (1999); and Davis *et al., Proc. Natl. Acad. Sci. USA, 96*:6541-6546 (1999), all of which are herein incorporated by reference in their entirety). The use of markers to isolate a genomic region of interest is often referred to as map based cloning, chromosome walking or positional cloning. Many of the *Arabidopsis thaliana* markers that have been used in map based cloning are anchored to genetic maps such as the Lister & Dean map (*See e.g.* genome-www3.stanford.edu/ cgi-bin/AtDB/ Riintromap).

Please replace the paragraph on page 2, line 26 to page 3, line 7 with the following:

Physical or partial physical maps of the *Arabidopsis thaliana* genome have also been reported (*See e.g.* genome-www3.stanford.edu/atdb_welcome.html). A physical map of *Arabidopsis thaliana*, Columbia based on a collection of bacterial artificial chromosomes (BACs) is available (Marra *et al., Nat. Genet., 22(3)*:265-270 (1999); Mozo *et al., Nat. Genet., 22(e)*:271-275 (1999), both of which are herein incorporated by reference in their entirety). An overlapping series of BACs representing the *Arabidopsis thaliana*, Columbia genome is available from AIMS, Arabidopsis Biological Resource Center, 309 B&Z Building, 1735 Neil Avenue, Columbus, OH 43210, USA.

Please replace the paragraph on page 8, line 24 to page 9, line 14 with the following:

The primers herein are selected to be "substantially" complementary to the different strands of each specific sequence to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands. Therefore, the primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the strand to be amplified to hybridize therewith and thereby form a template for synthesis of the extension product of the other primer. Computer generated searches using programs such as Primer 3 (www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), STSPipeline (www-genome.wi.mit.edu/cgi-bin/www-STS_Pipeline), or GeneUp (Pesole *et al., BioTechniques 25*:112-123 (1998), the entirety of which is herein incorporated by reference), for example, can be used to identify potential PCR primers. Exemplary primers include primers that are 18 to 50 bases long, where at least between 18 to 25 bases are identical or complementary to at least 18 to 25 bases of a segment of the template sequence.

Please replace the paragraph on page 17, line 25 to page 19, line 12 with the following:

SNPs can be characterized using any of a variety of methods. Such methods include the direct or indirect sequencing of the site, the use of restriction enzymes (Botstein *et al., Am. J. Hum. Genet. 32*:314-331 (1980), the entirety of which is herein incorporated by reference; Konieczny and Ausubel, *Plant J. 4*:403-410 (1993), the entirety of which is herein incorporated by reference), enzymatic and chemical mismatch assays (Myers *et al., Nature 313*:495-498 (1985), the entirety of which is herein incorporated by reference), allele-specific PCR (Newton *et al., Nucl. Acids Res. 17*:2503-2516 (1989), the entirety of which is herein incorporated by reference; Wu *et al., Proc. Natl. Acad. Sci. USA 86*:2757-2760 (1989), the entirety of which is herein incorporated by reference), ligase chain reaction (Barany, *Proc. Natl. Acad. Sci. USA 88*:189-193 (1991), the entirety of which is herein incorporated by reference), single-strand

conformation polymorphism analysis (Labrune *et al., Am. J. Hum. Genet. 48*:1115-1120 (1991), the entirety of which is herein incorporated by reference), single base primer extension (Kuppuswamy *et al., Proc. Natl. Acad. Sci. USA 88*:1143-1147 (1991), Goelet US 6,004,744; Goelet 5,888,819; all of which are herein incorporated by reference in their entirety), solid-phase ELISA-based oligonucleotide ligation assays (Nikiforov *et al., Nucl. Acids Res. 22*:4167-4175 (1994), dideoxy fingerprinting (Sarker *et al., Genomics 13*:441-443 (1992), the entirety of which is herein incorporated by reference), oligonucleotide fluorescence-quenching assays (Livak *et al., PCR Methods Appl. 4*:357-362 (1995a), the entirety of which is herein incorporated by reference), 5'-nuclease allele-specific hybridization TaqMan™ assay (Livak *et al., Nature Genet. 9*:341-342 (1995), the entirety of which is herein incorporated by reference), template-directed dye-terminator incorporation (TDI) assay (Chen and Kwok, *Nucl. Acids Res. 25*:347-353 (1997), the entirety of which is herein incorporated by reference), allele-specific molecular beacon assay (Tyagi *et al., Nature Biotech. 16*:49-53 (1998), the entirety of which is herein incorporated by reference), PinPoint assay (Haff and Smirnov, *Genome Res. 7*:378-388 (1997), the entirety of which is herein incorporated by reference), dCAPS analysis (Neff *et al., Plant J. 14*:387-392 (1998), the entirety of which is herein incorporated by reference), pyrosequencing (Ronaghi *et al, Analytical Biochemistry 267*:65-71 (1999); Ronaghi *et al* PCT application WO 98/13523; Nyren *et al* PCT application WO 98/28440, all of which are herein incorporated by reference in their entirety; www.pyrosequencing.com), using mass spectrometry *e.g.*, the Masscode ™ system (Howbert *et al* WO 99/05319; Howber *et al.* WO 97/27331, all of which are herein incorporated by reference in their entirety; www.rapigene.com; Becker *et al* PCT application WO 98/26095; Becker *et al* PCT application; WO 98/12355; Becker *et al.* PCT application WO 97/33000; Monforte *et al* US 5,965,363, all of which are herein incorporated by reference in their entirety), invasive cleavage of oligonucleotide probes (Lyamichev *et al Nature Biotechnology 17*:292-296, herein incorporated by reference in its entirety; www.twt.com), using high density oligonucleotide arrays (Hacia *et al. Nature Genetics 22*:164-167; herein incorporated by reference in its entirety; www.affymetrix.com).

B3 cont.

Please replace the paragraph on page 47, lines 16-25 with the following:

PHRED is used to call the bases from the sequence trace files www.mbt.washington.edu). PHRED uses Fourier methods to examine the four base traces in the region surrounding each point in the data set in order to predict a series of evenly spaced predicted locations. That is, it determines where the peaks would be centered if there are no compressions, dropouts, or other factors shifting the peaks from their "true" locations. Next, PHRED examines each trace to find the centers of the actual, or observed peaks and the areas of these peaks relative to their neighbors. The peaks are detected independently along each of the four traces so many peaks overlap. A dynamic programming algorithm is used to match the observed peaks detected in the second step with the predicted peak locations found in the first step.

B4

Please replace the paragraph on page 48, lines 3-12 with the following:

Contigs are assembled using PANGEA clustering tools (PANGEA SYSTEMS. INC) and PHRAP (www.mbt.washington.edu). PANGEA clustering tools are a series of scripts which group sequences (clusters) by comparing pairs of sequences for overlapping bases. The overlap is determined using the following high stringency parameters: word size = 8; window size = 60; and identity is 93%. Each of the clusters are then assembled using PHRAP. The final assembly output contains a collection of sequences including contigs, sequences representing the consensus sequence of overlapping clustered sequences, and singletons, sequences which are not present in any cluster of related sequences. Collectively, the contigs and singletons resulting from a DNA assembly are referred to as islands.

B5

Please replace the paragraph on page 49, line 14 to page 49, line 11 with the following:

INDELs are identified by aligning sequences from *Arabidopsis thaliana*, Columbia and *Arabidopsis thaliana*, Landsberg *erecta*. Finished BAC sequences derived from *Arabidopsis thaliana*, Columbia are obtained from GenBank (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide). Because the GenBank sequences are subject to change, the finished sequences of the *Arabidopsis thaliana*, Columbia BACs are included herein as SEQ ID NO: 1 through SEQ ID NO: 124. The sequence of each *Arabidopsis*

B6

thaliana, Columbia BAC is used as a query against a database of *Arabidopsis thaliana*, Landsberg *erecta* islands using the GAP2 program of the Analysis and Annotation Tool (AAT) for Finding Genes in Genomic Sequences which was developed by Xiaoqiu Huang at Michigan Tech University and is available at the web site genome.cs.mtu.edu/. See Huang, *et al.*, *Genomics 46*:37-45 (1997) and Huang, *Computer Applications in the Biosciences 10* 227-235 (1994), both of which are herein incorporated by reference in their entirety. The GAP2 program compares the query sequence with a cDNA database using a fast database search program and a rigorous alignment program. The database search program quickly identifies regions of the query sequence that are similar to a database sequence. Then the alignment program constructs an optimal alignment for each region and the database sequence. The output file of GAP2 is reviewed for insertions or deletions. Using alignments that are at least 96% identical (as reported by AAT), insertions and deletions are determined by looking for gaps of at least three bases, with three aligned bases on either side of the gap. To ensure that an insertion or deletion is derived from matched sequence, the 10bp region to either side of the gap is aligned and compared. To be considered an insertion or deletion, the adjacent aligned regions must be at least 90% identical (as reported by AAT). Insertions or deletions smaller than 100bp are considered candidate markers. INDELs identified by the method of this Example 2 are set forth in Table A and identified in the "method" column by reference to method 2. More particularly Table A identifies the location and nature of the polymorphism as follows.